Parameter Optimization of Infection Models

András Csernenszky, Gyula Kovács, Miklós Krész, and András Pluhár

Recently one of the main directions of datamining is the study and use of network data. Our research is concentrated on the network data about the links of big and medium size of companies that can be deduced from the bank transactions. A year ago the main goal for us was to develop an universal model to predict churn and bankruptcy. That work resulted in the implementation of the Domingos-Richardson cascade model for bankruptcy forecasting in 2009 September, which has since performed in the Bank quite well. The aim of our current research is to re-examine and redefine parameter optimization options and handle this issue in a more elaborate way. The performance of the method greatly depends on the estimation of the infection probabilities. In this research we get the better influence values assigned to the edges (which are representing the connections between two companies) that futher improves the infection predictions. The cornerstones for these improvements are the involvement on certain static variables and an appropriate optimization on the parameters that take their effects into account.

References

[1] P. Domingos, M. Richardson, Mining the Network Value of Costumers.*7th Intl. Conf. on Knowl-edge Discovery and Data Mining*, 2001.